

Assessing the Value of Structured Analytic Techniques in the U.S. Intelligence Community

Stephen Artner, Richard S. Girven, and James B. Bruce

Key findings

- The U.S. Intelligence Community does not systematically evaluate the effectiveness of structured analytic techniques, despite their increased use.
- One promising method of assessing these techniques would be to initiate qualitative reviews of their contribution in bodies of intelligence production on a variety of topics, in addition to interviews with authors, managers, and consumers.
- A RAND pilot study found that intelligence publications using these techniques generally addressed a broader range of potential outcomes and implications than did other analyses.
- Quantitative assessments correlating the use of structured techniques to measures of analytic quality, along with controlled experiments using these techniques, could provide a fuller picture of their contribution to intelligence analysis.

SUMMARY ■ The Intelligence Community (IC) is strongly emphasizing the use of structured analytic techniques (SATs) to promote rigorous analysis, lessen the risk of intelligence failure, and make analysts' reasoning more transparent to consumers.

So far, however, the IC has made little effort to assess whether SATs in general or specific SATs are improving the quality of analysis.

One primarily qualitative method of evaluating these techniques would be periodic in-depth reviews of IC production on a variety of topics to ascertain how frequently SATs are used, whether they advance the attainment of IC analytic quality standards, and which specific SATs are most effective.

RAND's preliminary SAT review of a limited sample of Central Intelligence Agency (CIA), Defense Intelligence Agency (DIA), and National Intelligence Council (NIC) publications finds that the minority of them employing SATs addressed a broader range of potential outcomes and implications than did other analyses, although the logic behind specific techniques was not always transparent.

More-comprehensive SAT reviews, along with interviews of analysts, managers, and consumers, could determine how effectively IC agencies are employing SATs and compile lessons learned on best practices and pitfalls of their use in intelligence analysis.

In addition, quantitative research could examine the extent to which the use or nonuse of SATs correlates with the quality of analytic output as measured by the Office of the Director of National Intelligence's (ODNI's) Office of Analytic Integrity and Standards and by individual agencies, while controlled experiments could test the use of SATs under conditions similar to those prevailing in the IC.

INTRODUCTION

In recent years, the IC has greatly increased its use of SATs to promote rigorous analysis, lessen the risk of intelligence failure, and make analysts' reasoning more transparent to consumers. So far, however, the IC has made little effort to assess the impact of SATs on the quality of intelligence analysis. This report outlines a research design for systematically reviewing the use of SATs in IC-wide finished production to assess their value in general and the utility of particular SATs for addressing various types of intelligence issues. Introducing a variant of the analytic line review, we propose SAT reviews on discrete bodies of IC work and present a pilot project on a limited sample of IC finished production. The pilot project was intended to illustrate how broader SAT reviews could draw conclusions regarding the overall value of SATs, the extent of their utilization in IC analysis, the suitability of particular SATs to specific problem sets, possible disadvantages of SATs, best practices in their use, and the need for additional analytic methodologies. The report also proposes several other approaches the IC could utilize to evaluate the effectiveness of SATs and optimize their use in finished production.

What Are Structured Analytic Techniques, and Why Do Analysts Use Them?

SATs are methods of organizing and stimulating thinking about intelligence problems. These methods aim to make the analytic process conscious and transparent, thus reducing the probability of errors caused by numerous cognitive biases that go unchallenged in unstructured and intuitive analysis. According to an unclassified U.S. government guide to analytic tradecraft,¹ SATs fall into three broad categories: *Diagnostic techniques* aim to make assumptions and logical arguments more transparent; *contrarian techniques* challenge current thinking; and *imaginative thinking techniques* encourage new perspectives, insights, and alternative scenarios.

Among the many SATs in use today, the most common are

- *structured brainstorming*—a group discussion process used for generating new ideas and concepts often used to kick off analysis of especially complex or controversial issues
- *key assumptions check*—a systematic effort to make explicit and question the assumptions that guide an analyst's interpretation of evidence and reasoning about any particular problem
- *analysis of competing hypotheses (ACH)*—the identification of a complete set of alternative hypotheses, the systematic

evaluation of each through the examination of evidence and data that applies to them all, and the selection of the most explanatory or best-fitting hypothesis (or hypotheses) by focusing on information that tends to disconfirm weaker hypotheses

- *indicators*—a preestablished set of observable phenomena that are periodically reviewed to help track events, spot emerging trends, and warn of unanticipated changes.

SATs provide analysts with clear, often step-by-step, guidance for conducting analysis of intelligence issues. By providing greater structure to the analytic process, they reduce subjectivity and add both rigor and transparency to analysis. A key part of reducing subjectivity in analysis requires identifying cognitive bias and reducing it. Chief among such biases often seen in intelligence analysis are

- *confirmation bias*—a tendency to search for or interpret information in ways that confirm preconceptions, preferences, and assumptions, while downplaying or discrediting alternative or less-agreeable explanations that tend not to confirm the preferred explanation or interpretation of events
- *mirror imaging*—an inclination to assume that foreign leaders would behave pretty much as we imagine our own leaders would behave in similar circumstances, especially when the stakes are high if major errors are made in “rational” decisionmaking
- *anchoring*—a tendency to “anchor” analysis in the first or earliest important piece of information considered, so that later changes in judgments are typically small and rarely stray far from the initial judgment
- *groupthink*—a usually subconscious preference for group consensus favoring agreement among group members and subtly discouraging alternative views and interpretations, which are often seen as efforts to disrupt the consensus the other members desire.

In addition to their significant debiasing capabilities, SATs also provide step-by-step guidance for designing analysis to address particular kinds of problems. They make analysis much more transparent, so that other analysts and customers can better understand how the judgments were reached.

SATs also facilitate group involvement in analysis because multiple analysts can use these techniques collaboratively on difficult problems and, in the process, help reduce potential biasing effects of individual analysts while highlighting the bases for analytic agreement and disagreement.

Additionally, through greater analytic transparency, SATs help analysts evaluate how much confidence they should have in particular judgments, and some can even help in estimating rough likelihoods that some event or outcome may or may not occur. Use of SATs also supports and helps to implement the important analytic and tradecraft standards issued by the Office of the Director of National Intelligence (ODNI) in Intelligence Community Directive (ICD) 203.2

Methodology

We interviewed analytic managers and analytic methodologists to determine the extent of SAT training and utilization, the emphasis individual agencies place on SAT use, the suitability of particular SATs to specific problem sets, possible disadvantages of SATs, best practices in their use, and the need for additional analytic methodologies. We also held an analyst focus group to discuss similar issues from the perspective of analysts who choose or are required to use SATs. We then examined a small sample of IC finished intelligence products to ascertain how many of them used SATs, how particular SATs did or did not contribute to the analytic quality of the paper, and how papers that did not use SATs compared with those that did on IC tradecraft standards. We discuss the specifics of the study sample later, under “Structured Analytic Technique Review: A Pilot Study.”

THE INCREASED PROMINENCE OF STRUCTURED ANALYTIC TECHNIQUES IN THE INTELLIGENCE COMMUNITY

The IC has increased its use of SATs in the years following the intelligence failures on Iraqi weapons of mass destruction. The late analytic tradecraft expert Jack Davis has noted that the CIA vigorously promotes alternative analysis formats and structured analysis, including the systematic generation and rigorous review of alternative hypotheses.³ A senior DIA officer has stated that his agency is systematically training all new analysts in four key structured techniques and is using these and other SATs in a variety of products.⁴ U.S. Marine Corps Intelligence, working with the RAND Corporation, has developed a set of 28 structured models, approaches, and techniques that apply structured analysis to tactical battlefield situations.⁵ ODNI promotes IC-wide training for analysts in critical thinking and SATs,⁶ and the 2014 National Intelligence Strategy calls on the IC to reinforce analytic methods that challenge long-standing assumptions and encourage new perspectives.⁷

Analytic tradecraft experts note that SATs have gained in popularity in part because they specifically address cognitive pitfalls that have been associated with recurrent intelligence failures. For example, various SATs address mindset bias; “satisficing” (settling on the first plausible hypothesis); and the need for greater imagination, the importance of validating evidence and sources, and the unreliability of single-point predictions. The explicit use of SATs in finished products can make analysts’ logic more transparent to readers who want to know how analytic conclusions were reached, allowing readers to use different assumptions and information to challenge these conclusions. SATs also capitalize on the benefits of collaboration among groups consisting of analysts with different backgrounds and perspectives.⁸

DEBATING THE VALUE OF STRUCTURED ANALYTIC TECHNIQUES

Intelligence practitioners cite logical reasons to assume that SATs are beneficial. Heuer and Pherson noted that, “if a structured analytic technique is specifically designed to mitigate or avoid one of the proven problems in human thought processes, and if the technique appears to be successful in doing so, that technique can be said to have face validity.”⁹ The investment of considerable time, expense, and personnel in the promotion of SATs suggests that the IC agencies’ experience with the techniques has generally been positive. According to our interviews, some regional IC offices use SATs widely, especially on hard issues for senior customers, reportedly with good results.

Other observers argue that structured techniques are impractical to use in the IC or are inferior to unaided expert judgment. As noted almost universally in our unclassified interviews, many analysts consider SATs too time consuming, especially for short-deadline current intelligence products, although advocates reply that the use of SATs early in a project can save time later in editing and coordination. Some critics have contended that SATs eliminate the use of insights from expert intuition, lack the flexibility to handle rapidly changing conditions, or cannot account for all the variables present in complex problems.¹⁰ One critique of the ACH method argues that it is cumbersome to use in practice, requires an inordinately large number of judgments to be based on unarticulated assumptions, and encourages analysts to overlook important contextual information.¹¹

More important, even if SATs in general are valuable, they can be misapplied in individual cases. Proponents of SATs

The IC has increased its use of SATs in the years following the intelligence failures on Iraqi weapons of mass destruction.

describe them as tools to encourage collaborative critical thinking and make analysts' logic more transparent to themselves and to consumers.¹² Like any tool, SATs depend on the analytic skills and substantive expertise of their users to be effective. They are unlikely to benefit analysis if they become mechanical processes or box-checking exercises rather than aids to imaginative thinking.¹³

THE LACK OF STRUCTURED ANALYTIC TECHNIQUE EVALUATION

SATs have received little formal evaluation so far, despite their increased use in the IC.¹⁴ Heuer noted in 2009 that no element of the IC had the charter or resources to systematically test SATs.¹⁵ Our research and interviews with a range of IC managers, analysts, and methodologists in 2014 indicate that none of their agencies is routinely measuring the use of SATs or how they affect the quality of analysis. The few relevant studies are dated, produced mixed results, and did not fully replicate conditions in the IC:

- A study published by the Joint Military Intelligence College in 2000 found that structured hypothesis testing markedly improved analytic accuracy in one of two controlled experiments among analysts in theater Joint Intelligence Centers but had little effect in the other.¹⁶
- A controlled experiment conducted by the Mitre Corporation in 2004 found that the use of ACH—one of the most important SATs—reduced confirmation bias only among participants who lacked a professional intelligence background. The experiment measured the effect of ACH use on individual participants but did not involve group interactions.¹⁷

- In a book published in 2005, Philip Tetlock described two experiments showing that scenario development reduced the accuracy of predictions.¹⁸ However, as Heuer and Pherson note, the IC does not use scenarios as a predictive tool but as a way to outline several futures that policymakers may encounter.¹⁹
- Academic studies have found that the use of the devil's advocate technique does not necessarily promote genuine reexamination of assumptions and, in some cases, heightens confidence in preferred hypotheses.²⁰
- A 2009 journal article by Sundri Khalsa, a Marine Corps intelligence expert, cites numerous academic studies suggesting that systemic processes combining structured techniques with intuitive judgments improve outcomes across a broad range of research fields including natural sciences, medicine, oil exploration, and psychology. Few of the cited studies, however, focused specifically on the use of SATs in the IC.²¹

The lack of systematic SAT assessment applies not only to SATs as a group but to individual techniques. Little or no evidence verifies whether particular SATs are effective at mitigating the cognitive pitfalls they are designed to address, dealing with particular kinds of substantive intelligence issues, or promoting effective group collaboration. Among the open questions regarding SATs are the following:

- How widely are SATs used in the various IC agencies? How frequently are they used in various types of intelligence products?
- Which SATs are most suitable for which kinds of problems?
- What are the most important best practices and pitfalls in the use of SATs?
- Do intelligence consumers think SATs improve analysis?
- Do SATs lead analysts to examine hypotheses and outcomes they otherwise would not have considered?
- To what extent do SATs help overcome mindset bias, satisficing, and other cognitive pitfalls that plague analysis?
- Can one demonstrate whether SATs have improved warning or the accuracy of analytic judgments?
- Are there significant opportunity costs and negative unintended consequences in the use of SATs?

DIFFICULTIES IN VALIDATING STRUCTURED ANALYTIC TECHNIQUES

The difficulty of evaluating analysis in general, and SATs in particular, helps explain why they have not been assessed in greater detail.²² Our interviews and other available unclassified information suggest that most IC agencies do not systematically track how often their products are using SATs and how SATs are affecting the quality of analysis. Individual agencies and ODNI regularly evaluate a sample of finished production for analytic quality, but these efforts do not explicitly examine the use and effectiveness of SATs. These evaluations also generally look at individual products rather than at bodies of work on a given topic, which might allow an assessment of how SATs or other approaches shape analysis over time. Finally, for a variety of reasons, the evaluation programs do not attempt to assess the accuracy of analysis.²³ As a result of these factors, few statistical or qualitative data are available on the use of SATs in the IC.

Several considerations would complicate the effort to introduce systematic evaluation of SATs:

- It is not always clear whether the author or authors of a given finished product made use of SATs unless this is stated in a scope note. Some SATs, such as structured scenarios, would be obvious in the final paper, but others, such as facilitated brainstorming or assumptions checks, might not be apparent.

Our interviews and other available unclassified information suggest that most IC agencies do not systematically track how often their products are using SATs and how SATs are affecting the quality of analysis.

- A reviewer would need to make a reasoned judgment about how and whether a particular SAT contributed to a product's key judgments and analytic value, including whether the SAT enhanced quality beyond what would have otherwise been the case. In some cases, this may be apparent, but in others, the same conclusions might have been reached without the use of SATs.
- Interviews with analysts and managers can help address both sets of questions, but this may not be practical for evaluating a large body of IC products.

The reaction of policy consumers may also be difficult to gauge even if they consent to interviews or surveys. Some intelligence consumers may express an opinion on SATs, but one cannot assume that these individuals have strong feelings on the topic; their reactions to finished intelligence may depend on factors other than its analytic quality. For example, whether intelligence provides useful factual information or ammunition in intra-administration or partisan policy disputes may be at least equally important to some readers.

STEPS TOWARD SYSTEMATIC EVALUATION OF STRUCTURED ANALYTIC TECHNIQUES

The difficulties of assessing SATs suggest that using several different approaches is more promising than limiting the evaluation to a single method. The RAND pilot project documented here explored two options—structured interviews of analytic practitioners and qualitative reviews of IC production—in addition to suggesting several other possibilities for future evaluation. Our intent is not to offer definitive conclusions about the value of SATs but to propose ways that the IC can begin more-systematic evaluation of these techniques to assess their utility and optimize their use in analysis.

Interviewing practitioners and methodologists is one essential component of any evaluation effort. Those who use (or choose not to use) SATs in their daily work can offer invaluable insights into the value of such techniques for quality of analysis. In some cases, an individual might have a bias—current intelligence analysts who operate under tight production deadlines might have a different mindset from do methodologists whose jobs involve facilitating the use of SATs. But such practitioners and their supervisors are the group for which SATs have been devised, who have the most direct experience with their

use, and who can therefore offer informed comments on their strengths and drawbacks.

A second approach is to examine the use of structured techniques in samples of IC production in what we propose calling *SAT reviews*. These would be a variant of the *analytic line review*, an art that is gaining usage in some parts of the IC to allow a qualitative evaluation of a body of finished products on a particular topic over a given period:

- An analytic line review carefully examines each published item in such a body of production to determine whether the IC, or a particular agency, had a discernable primary analytic line on the topic throughout the period under review; whether the line was consistent or shifted over time; and, if it shifted, whether the reasons for the changes were explained to the reader. An analytic line review can also attempt to assess the overall utility of the analysis to policymakers, its analytic quality as defined by ODNI or agency tradecraft standards, and—if written after a major discontinuity or surprise—whether the analysis provided effecting warning.
- SAT reviews could take a broadly similar approach but focus on the use of analytic techniques in a body of work more than on the analytic message. The purpose would be to ascertain the extent to which a body of IC production employed SATs; which ones were used; whether and how they affected key analytic judgments; whether they contributed to analytic quality, as defined by IC standards; and, when possible, whether they promoted analytic accuracy and warning.
- SAT reviews would be qualitative rather than quantitative. However, if the IC undertook numerous such reviews over an extended period, useful metrics could emerge on the extent to which various SATs are used across a swath of IC production and on their contribution to analytic quality.

Later, for future consideration, we will briefly describe some additional approaches to evaluating SATs.

We are proposing a multipart strategy to assess SATs because each individual approach has drawbacks and limitations. For example, interviews and SAT reviews each include subjective elements that careful research design can limit but not eliminate entirely. Replication of such efforts over a longer period and with larger sample sizes, however, could increase confidence in key conclusions.

Those conducting SAT reviews may not know whether SATs were employed unless this is noted in a scope note or is

obvious from the text. However, in larger follow-on projects to this one, reviewers could contact authors of the pieces to request this information and to solicit authors' own assessments of whether and how SATs affected their analysis.

STRUCTURED ANALYTIC TECHNIQUE REVIEW: A PILOT STUDY

We examined a small sample of finished intelligence products to ascertain how many used SATs, how particular SATs did or did not contribute to analytic quality, and how papers that did not use SATs compared with those that did on IC tradecraft standards. The sample includes a set of CIA intelligence assessments (IAs), NIC analytic products published in July 2014, and a random set of DIA and CIA papers published in 2013 on several selected intelligence issues.

The conclusions from this sample are not meant to be applicable to IC products in general. Rather, this study is designed to illustrate how a much larger-scale SAT review of finished products could lead to firmer and more granular judgments about the use and value of SATs. For example, a larger study could examine the value of additional types of SATs; if that study were extended over time, multiple agencies, and a variety of substantive issues, could begin to accumulate meaningful data on the IC's use of SATs.

CIA INTELLIGENCE ASSESSMENTS

RAND reviewed 29 CIA IAs posted on the World Intelligence Review electronic (WIRE) site during a two-week period in July 2014. The IA is the longest form of current intelligence the CIA publishes. We chose to examine CIA IAs in this part of our study because that agency puts particular weight on structured techniques, according to Richards Heuer and several methodologists with whom we spoke.²⁴ We also assumed that SATs are most likely to appear in longer papers, based on the IC experience of RAND researchers and on the frequently made comment in our interviews that many analysts consider SATs too time consuming to use in shorter current intelligence products.

Of the CIA IAs in our sample, 23 showed no evidence of using SATs. It is possible that analysts used ACH, assumption checks, structured brainstorming, or other structured techniques during the conceptualization and research phases of the IAs without saying so in the final text. However, this seems

We examined a small sample of finished intelligence products to ascertain how many used SATs, how particular SATs did or did not contribute to analytic quality, and how papers that did not use SATs compared with those that did on IC tradecraft standards.

unlikely in most cases because almost all the papers included scope notes that discuss the projects' research design. In addition, we did not find any indirect evidence of SAT use in these IAs, such as detailed discussion of alternative hypotheses not explicitly linked to the use of an SAT.

Six of the IAs in our sample made explicit and extensive use of SATs. Three of these IAs used structured alternative scenarios developed through facilitated brainstorming; one was constructed around an assumptions check; one presented a Team A/Team B analysis; and one created a detailed set of indicators to evaluate future developments.

Many of the IAs that did not employ SATs offered incisive and policy-relevant analysis, in our judgment. For example, some of these offered valuable information and insights on countries or topics of high importance to U.S. national security, discussed foreign demographic trends that could affect U.S. interests, and provided scientific expertise that is highly relevant to arms control negotiations. Some of the IAs, in our view, did an outstanding job of forward-leaning analysis focused on threats to U.S. interests. Some also highlighted opportunities for U.S. policymakers to affect outcomes or mitigate negative consequences for the United States.

A small number of the IAs that did not employ SATs showed tradecraft problems, in our opinion. In one paper, the key judgment is expressed as a tautology that is not falsifiable by events. In another case, the title did not reflect the analytic thrust of the assessment, and its key message was therefore somewhat unclear. Although the use of SATs would not necessarily have eliminated these problems, it might have promoted more group collaboration that could have detected and addressed them.

The more-general pattern in the IAs that did not use SATs was a focus on single-point outcomes assuming continuity. These IAs offered a consensus view of current developments,

did not state or question the assumptions underlying that consensus view, and posited that future developments would reflect either the continuation of the status quo or the extrapolation of current trends. Although these assessments offered valuable intelligence reflecting deep expertise, we believe that, in some cases, they would have been even more useful if they had made at least a limited attempt to consider alternative interpretations to the mainline view.

Use of Indicators

One IA developed a set of indicators to use in tracking the spread of a certain form of instability. The study drew on statistical modeling and a range of academic case studies to offer a framework for analysis while avoiding single-point predictions. The assessment builds on the indicator methodology that has traditionally been a key part of warning analysis. In this case, the IA adapted the technique to apply it to a set of problems that is more marked by nonlinearity and complexity than are traditional warning issues of hostile action by a unitary decisionmaker.

One drawback of indicator analysis, particularly when used on nonlinear issues, is that it is based on historical precedents that may or may not signal future discontinuities. The IA offers a structured guide, however, to determining when historical analogies are valid comparisons and when they are not.

The IA responded to a specific request from senior policymakers. Its approach would not, in our view, be suitable for most finished intelligence but serves as a guide for analysts and other subject-matter experts.

Key Assumptions Check

Another IA on a high-priority topic rested in large part on a check of key assumptions conducted to challenge mainstream thinking. After identifying four key assumptions underlying a widely shared key judgment, the IA assessed evidence supporting or contradicting each assumption, noted information gaps, examined implications if the assumption was mistaken, and checked for indications that the assumption was previously correct but is so no longer. An appendix includes detailed discussion of the four key assumptions, noting the degree of confidence in each—including dissenting views—and judging how critical each was to bottom-line conclusions.

In our judgment, this IA offers effective forward-looking analysis by identifying probable key drivers of the interactions between key actors in what is likely to be a volatile and unpredictable situation. The assumption check approach also helped the analysts examine factors that could change the key drivers and allowed discussion of a range of plausible outcomes. This approach probably would not have lent itself to a nonspecialist senior consumer who wanted a single projection of a most likely outcome. This IA is probably of greatest use to mid- and working-level readers who are deeply immersed in the subject matter, familiar with the key factors and actors at play, and focused on developing U.S. strategies to deal with a range of potential outcomes.

Team A/Team B

One IA used a Team A/Team B exercise to examine potential political outcomes in a country of interest. The technique allowed the full development of two competing hypotheses and avoided the lowest-common-denominator analysis that characterizes some fully coordinated IC products. Several steps might have made the analysis even more useful, in our view.

The IA might have indicated whether the exercise represented a genuine difference of opinion among analysts or merely an analytic technique. If the former, it would be of interest to know whether there were majority and minority opinions or whether particular groups of analysts—by discipline, substantive focus, or office—leaned toward one thesis or the other.

It also would be useful to establish specific indicators that would signal which outcome is becoming more likely. In the absence of such indicators—and of explicit implications for the United States of each scenario—this IA appears useful more as an internal analytic exercise than as a product for senior policymakers.

Facilitated Brainstorming and Alternative Scenarios

Another IA used facilitated brainstorming to anticipate the prospects for democracy, stability, and shared national identity in a foreign country. This approach promoted a detailed analytic discussion of a most likely trajectory by isolating and analyzing key drivers. In addition, the assessment noted system shocks that could alter that current trajectory and cited two alternative sources of national identity that would present radically different outcomes from the country's current course. A brief section on implications includes both mainline and alternative scenarios.

One useful aspect of this IA was that it identified both a most likely outcome and two alternatives. Many non-U.S. government experts in scenarios argue they should not include probability statements because these divert attention from alternatives that are also highly plausible, even if they seem less likely at the moment. This position is debated within the IC, with some analysts agreeing and others believing that policymakers are best served by attaching probability statements to scenarios. In this case, CIA analysts apparently agreed that one outcome was clearly the most likely, and policymakers would not have been well served by pretending that all three outcomes were equally probable. At the same time, the use of alternative scenarios took account of the intrinsic limitations of long-range predictions, especially on a subject as broad as that discussed in the IA, and prepared policymakers for different sets of implications in case the mainline projection turns out to be incorrect.

Another IA used alternative scenarios based in part on structured brainstorming to discuss prospects for an ideological movement in an important region. The study began with a detailed analysis of actors and factors promoting and opposing the spread of that ideology. Based on this analysis, the outlook section begins with a mainline forecast positing only slow growth in support for the ideology. The assessment then discussed four sets of alternative future developments that could radically intensify or undermine the spread of the ideology and affect the future scope of its influence.

Like the IA on scenarios described earlier, this one combined a mainline forecast that is likely to be of most direct interest to policymakers and a discussion of factors that could fundamentally alter this projection. In this case, the analysis focused more on key drivers of alternative outcomes than on future states. The consideration of alternative scenarios also facilitated discussion of steps U.S. officials might consider to mitigate the consequences of the most negative scenarios.

One IA revisited scenarios developed in a 2013 paper on the implications of international trends for political stability in a key country. This one offered detailed discussion of how the scenarios could play out and prompt varying reactions from different parts of the country in question. It also discussed implications for the United States. The analysis might have been even sharper in two areas:

Although the IA referred in many places to how various scenarios would affect the country's stability, the discussion focused only on the degrees of difficulty the country would have in dealing with these scenarios, rather than on factors that would seriously threaten the regime's survival.

The IA states that two drivers will be decisive but does not explain why they were selected over others, perhaps because this was discussed in the earlier IA.

NATIONAL INTELLIGENCE COUNCIL ANALYTIC PRODUCTS

We also examined NIC analytic products posted on Intelink in July 2014. Our sample excluded short memos of fewer than five pages, conference reports, and a contract paper. The remaining documents included one National Intelligence Estimate, one NIC report, five IC assessments, and seven NIC memos. Of these 14 products, ten made no explicit use of SATs; two were centered on SATs; and two others mentioned them in a box or appendix. As with the CIA products we reviewed, our NIC sample was too small to allow definitive conclusions but is intended to suggest preliminary hypotheses for further evaluation.

One NIC publication that used SATs was based on a scenario exercise that examined trajectories for a country's economy based on exogenous political factors. It also included a box examining scenarios for political stability or instability. The item primarily outlined mainline projections for economic growth and political stability, but the discussion reflected nuances from the consideration of other scenarios. In addition, an appendix discussed the drivers and implications of three alternative scenarios.

Another document was based on a multiagency "what if" exercise that posited a negative event occurring and examined its potential causes and implications. The exercise examined potential catalysts for the event; factors that could affect its timing; and potential specific actors, targets, and methods. The item did not assess the likelihood of the event, in line with the "what if" methodology, but provided strong warning of its

potential and a detailed discussion of various ways by which it could come about.

In addition to the two documents centered on SATs, another included a short box on alternative scenarios. The scenarios do not directly contradict the item's key judgment, even though the authors express only low to moderate confidence in that conclusion. Another document includes a list of standing indicators in an appendix.

As in the case of CIA IAs, the NIC products that did not employ SATs addressed high-priority topics and demonstrated deep expertise. The NIC products also drew upon analytic insights from across the IC. Some of the documents that did not use SATs did briefly discuss wildcards or alternative outcomes. One of them explicitly discussed alternative scenarios but did not explain how they were arrived at.

In one case, a NIC memo relied on a crucial assumption that was open to challenge, in our opinion. We think the memo would have benefited from an explicit assumption check that weighed evidence and logic for and against the assumption and from at least a brief "what if" analysis examining the implications if the assumption turns out to be incorrect.

Many NIC documents list key assumptions in an opening section but do not offer alternative hypotheses. NIC publications also routinely address confidence levels in key judgments as directed in ODNI analytic tradecraft guidelines. Particularly in cases of low or moderate confidence, greater use of SATs would allow both more systematic scrutiny of the evidence on which judgments are based and the consideration of alternative theories, explanations, or scenarios. In some cases, a devil's advocacy or Team A/Team B exercise could be useful for challenging the basis of judgments in which the IC has high confidence, particularly to counter potential groupthink and make sure that other points of view are receiving consideration.

MULTITOPIC SAMPLE

We also reviewed a random sample of 20 DIA and CIA documents published in 2013 covering four selected intelligence issues. Of these, eight showed evidence of using at least one SAT: Five used alternative scenarios; one drew from a facilitated interagency brainstorming session; one included an indicators-based matrix as an appendix; and one used structured historical analogies.

All the documents showed considerable expertise and contributed to the U.S. government's understanding of high-

priority foreign policy challenges, in our opinion. Many drew on unique information that would not have been available in open media. Most also specifically addressed U.S. policy implications.

The documents that did not employ SATs did a strong job of reporting factual information. Some of these publications also provided forward-leaning analysis that addressed how recent and current developments could affect U.S. interests. The outlook sections of these products assumed that current trends would continue and did not discuss drivers or wildcards that could produce a different trajectory.

Those NIC products that discussed alternative scenarios avoided continuity-based single-point predictions. In most cases, these publications presented a mainline outlook based on the extrapolation of current trends but also discussed factors that could change the current trajectory and produce fundamentally different outcomes.

Some of the scenarios in this sample used innovative techniques:

One included a graphic based on the traditional two-by-two matrix showing combinations of two key drivers but superimposed three concrete scenarios based on specific foreign government actions at points along one of the axes. The graphic also indicated a range of uncertainty about how a second foreign government would respond, which is indicated as a range along the second axis.

Another presented scenarios from the viewpoint of the foreign government, which helped to anticipate its potential reactions under various outcomes. The text did not make clear whether the IC had reporting indicating that the foreign government was envisaging the particular scenarios, as some of the language seemed to imply, or whether they were developed by IC analysts attempting to put themselves in the place of the foreign leaders.

In most cases, the papers that offered alternative scenarios could have done more to explain the underlying logic, in our assessment. Most of the scenario pieces in this sample did not explain how the scenarios had been developed. It would have been useful to know whether they were based on a structured methodology or on expert intuition. In the latter case, the reader would benefit from knowing what kinds of experts had been consulted—for example, only the authors of the paper, a broader range of analysts from one or more IC agencies, or outside experts. In addition, most did not provide concrete indicators that would signal which scenario (or combination of scenarios) was becoming more likely over time.

The use of SATs in a minority of CIA, DIA, and NIC documents provided more-sophisticated analysis than would have been possible with straight-line projections.

One document used structured analogies to draw lessons for a potential negotiation process. The study employed 29 historical case studies and particularly examined five cases that produced successful negotiations, in addition to drawing on interviews, conferences, and IC reporting. One useful component was a detailed matrix listing confidence-building measures that helped produce agreements in previous situations and noted their potential applicability to a present conflict. In our opinion, the document made careful and proper use of the analogies, highlighting both similarities and differences with current situations. It stressed underlying factors in past agreements that might apply to the present but avoided identifying any one case study as a model.

One study included an appendix based on a comprehensive set of indicators. This helped to inform the analysis in the main text and, by implication, offered guidelines for tracking future developments. The scope note of another document stated that it had used a structured interagency brainstorming session to delineate implications for the United States. The exact contribution of brainstorming to the analysis is not clear because the section on implications also drew from a range of IC reporting and from a study of the foreign country's past behavior.

CONTRIBUTIONS AND LIMITATIONS OF STRUCTURED ANALYTIC TECHNIQUES

We believe the use of SATs in a minority of CIA, DIA, and NIC documents provided more-sophisticated analysis than

The use of SATs, such as indicators and alternative scenarios, in forward-looking analysis is most compatible with an understanding of the mission of intelligence as helping policymakers understand, cope with, and influence ongoing developments, rather than offering definitive answers or predictions.

would have been possible with straight-line projections assuming continuity. These publications examined a broader range of potential outcomes and thereby helped consumers prepare to deal with implications should mainline projections prove inaccurate.

The utility attributed to SATs depends in part, however, on how one defines the purpose of intelligence in general and specific intelligence products in particular. The use of SATs, such as indicators and alternative scenarios, in forward-looking analysis is most compatible with an understanding of the mission of intelligence as helping policymakers understand, cope with, and influence ongoing developments, rather than offering definitive answers or predictions. This model of intelligence is comparable to the role of football scouts who try to help a coach assess, counter, and defeat the opposing team, rather than predicting who will win the game.

Each document in this sample that used SATs focused primarily on intelligence “mysteries”—contingent future developments, rather than current factual inquiries. Some consumers

may prefer precise predictions over the multiple trajectories outlined in these IAs. The documents that used alternative scenarios, however, went some way in this direction by identifying which outcomes were mainline projections and which were less likely alternatives. These pieces dealt with such broad strategic issues that we suspect many readers would have questioned the value or validity of discussing only a single outcome.

This sample did not include documents that used SATs to focus on intelligence “puzzles”—factual questions about the past or present that are definitively answerable given sufficient evidence. Some of the papers in the sample did examine such issues. In these cases, SATs that focus on scenarios and warning indicators probably would not have been appropriate. However, the use of diagnostic techniques, such as assumption checks, might have made the analysts’ reasoning process more transparent and allowed explicit statements of degrees of confidence in key judgments.

STRUCTURED ANALYTIC TECHNIQUES AND INTELLIGENCE COMMUNITY ANALYTIC TRADECRAFT STANDARDS

Our review of IC products is based on the standards of analytic tradecraft enunciated in ICD 203. Our review of a limited sample of IC products suggests that some high-quality finished intelligence does not explicitly utilize SATs but that, when used properly, they can often help analysts fulfill IC analytic standards.

In general, documents using SATs and those not using SATs both did a mixed but generally successful job of addressing IC tradecraft standards. Almost all the documents we reviewed demonstrated relevance to U.S. national security interests, although few consistently described the quality and reliability of underlying sources. NIC documents routinely expressed levels of confidence in key judgments, while those from CIA did so inconsistently, whether or not they employed SATs.

In our view, documents using SATs generally had an edge on the criteria of distinguishing between assumptions and judgments and of using logical argumentation. Some that did not use SATs did a good job of distinguishing assumptions from judgments, but this distinction was particularly clear in the small minority of documents that explicitly used assumption checks. Scenario pieces stood out in their use of logical argumentation by explicitly identifying key drivers and noting how different combinations of them could lead to varying trajec-

ries. In general, the use of SATs made logical inferences more transparent than would have been the case in documents based on intuitive analysis.

Documents using SATs had a clear advantage on the criterion “incorporates alternative analysis where appropriate.” In principle, those without SATs could have offered alternative hypotheses derived from intuitive, rather than structured, methods, but we did not observe this in the items we reviewed. In contrast, documents using SATs considered alternative assumptions or outcomes that constituted alternatives to main-line judgments.

Even some of the documents we reviewed that incorporated SATs did not fully succeed in fulfilling the injunction to “explain how alternatives are linked to key assumptions and/or assess the probability of each alternative”—part of the fuller description of alternative analysis in ICD 203. Neither did they all “include discussion of key indicators that, if detected, would help clarify which alternative hypothesis, viewpoint, or outcome is more likely or becoming more likely.” The scenario pieces did include some discussion of indicators, but these descriptions were sometimes so general that they probably would be difficult to detect or measure in practice.

Our review did not consider criterion 7 (consistency of analysis over time) because of the limited time frame and number of documents in our sample. This criterion would be a major focus of analytic line reviews looking at a considerable body of IC work over an extended period.

Our review also did not examine criterion 8 (accuracy), but this could be a focus for selected future SAT reviews. Observers have pointed out the difficulty of assigning overall accuracy scores to IC production: Among other things, this requires that the reviewer possess or have access to subject-matter expertise and know the correct answer to an intelligence question at the time of evaluation. Unless structured carefully, accuracy studies also run the risk of rewarding easy calls and penalizing analysis that offers judgments in the face of limited or contradictory information.²⁵ Nonetheless, under some conditions, SAT reviews could address the issue of analytic accuracy and the contribution to it, if any, that analytic techniques could make:

- An ideal case study on an intelligence *puzzle*—a factual question whose answer is, in principle, knowable—would select an issue on which the IC lacked definitive knowledge at the time of publication but subsequently obtained reliable reporting. For example, analysts may offer judgments about the properties of a secret foreign weapon system about which they later obtain reliable information because

IC Tradecraft Standards

ICD 203 (2007), states that IC finished intelligence

- Properly describes quality and reliability of underlying sources
- Properly caveats and expresses uncertainties or confidence in analytic judgments
- Properly distinguishes between underlying intelligence and analysts’ assumptions and judgments
- Incorporates alternative analysis where appropriate
- Demonstrates relevance to U.S. national security
- Uses logical argumentation
- Exhibits consistency of analysis over time, or highlights changes and explains rationale.

the system is used in battle; official announcements or leaks from the foreign government describe its qualities; or high-quality clandestine reporting on it becomes available.

- A case study on an intelligence *mystery*—a contingent development in the future that cannot be known with certainty at the time of writing—could focus on a major event, such as one country attacking another or a regime-changing coup, to examine the extent to which intelligence provided effective warning.
- In either type of case study, SAT reviews would examine not only the accuracy of key judgments but the contribution of analytic techniques in reaching the conclusions. Such studies could investigate the extent to which SATs helped analysts identify the correct answer when it ran counter to conventional wisdom. These reviews also could examine whether SATs added value in suggesting hypotheses, assumptions, and outcomes that went beyond summarizing hard evidence and pointed to what turned out to be accurate conclusions.

STRUCTURED ANALYTIC TECHNIQUE BEST PRACTICES AND PITFALLS

One major area for further research is best practices and pitfalls of SATs. Insights in this area would be most likely to come from a combination of broader SAT reviews and structured interviews with authors and methodologists.

Our pilot study did suggest several best practices, particularly in documents containing alternative scenarios, the SAT employed most frequently in our sample. One area for improvement in some of the documents was, in our opinion, greater transparency about the reasoning behind the scenarios:

- In particular, when intelligence analyses posit key drivers, it would be useful to know how they were selected and why they were considered more important than other potential drivers.
- In addition, it would be helpful to explain—perhaps in a box or appendix—the methodology that was used to construct the scenarios. These steps would help maximize the value of SATs in making clear to readers how analysts reached their key judgments.
- Scenario papers also are most useful when they include concrete and observable indicators that signal which outcome is becoming more likely, rather than broad generalities that are difficult to measure.

A key challenge for the scenario pieces, in our judgment, is tracking their findings over time. The papers will have only limited analytic value if they are one-time efforts, rather than baseline studies that will inform analysis on an ongoing basis. As noted earlier, one IA in our sample reassessed the drivers and implications of scenarios that had been delineated in a baseline study published in 2013.

More generally, when analysts produce documents constructed entirely around an SAT, we would suggest carefully considering whether these products are suited for senior consumers. In many cases, we doubt that high-level policy officials

will have the time or inclination to retrace the logic leading to several opposing outcomes if the analysis does not present a clear bottom line. Such documents can have high analytic value for analysts—and perhaps for working-level officials in policy agencies—in promoting the consideration of multiple hypotheses, drivers, and outcomes. The IC could consider establishing a line of internal products highlighting SATs to promote creative and divergent thinking that would not necessarily be disseminated to traditional senior customers. We believe analysts should be rewarded for producing such products as much as for publishing traditional finished intelligence, even if their dissemination is more limited.

FUTURE RESEARCH ON STRUCTURED ANALYTIC TECHNIQUES

As mentioned, this pilot study is intended only to illustrate the feasibility of broader qualitative work on SATs, rather than to offer definitive conclusions. A much broader study could involve larger-scale SAT reviews focusing on more-diverse bodies of work from a larger range of IC agencies:

- Follow-on work could focus on intelligence production on one or more issues, such as key countries of concern to the United States, emerging terrorist groups, aspects of weapons proliferation, or economic issues with strategic implications.
- Broader samples could cover a longer period than those used in this study and could include work from all IC agencies that produce finished intelligence on the topics being covered.
- Extended samples also could examine a broader array of products than the longer documents included in our pilot study and could attempt to ascertain the extent to which analysis based on SATs is used in briefings.
- Such larger-scale SAT reviews would allow the compilation of data indicating the extent to which SATs are used across a broad range of IC production and the specific techniques that are used most frequently.

More important, larger-scale SAT reviews would promote insights on ways to optimize the value of SATs in IC analysis. For example, such research would allow examination of the types of SATs that are most appropriate to particular kinds of intelligence problems:

When analysts produce documents constructed entirely around an SAT, we would suggest carefully considering whether these products are suited for senior consumers.

- Several different typologies could categorize the problems. One classification, used by Greg Treverton, sorts problems according to their degree of conceptual uncertainty: *Puzzles* involve factual questions; *mysteries* concern contingent developments, such as foreign leaders' decisions to launch military attacks or future election outcomes; and *complexities* relate to nonlinear, largely unpredictable mutual interactions among numerous interdependent actors and factors in an adaptive system, such as the world economy.²⁶
- Other typologies might sort intelligence issues by their substantive field (such as military, political, economic, leadership, targeting), by the type of collection that is most readily available on them (human, signals, measurement and signals, and open-source intelligence), or by the format analysts are using to address them (current intelligence, short intelligence memoranda, longer assessments, National Intelligence Estimates).
- Study of a broader swath of IC products, combined with interviews with practitioners and consumers, could ascertain the advantages and potential disadvantages of using particular kinds of techniques for various categories of problems. Such research also might highlight the need to further develop exiting SATs or to create new ones.

Broader research on SATs also would allow a greater focus on best practices and pitfalls in their use. As part of such an effort, interviews with the authors of specific products and an independent evaluation of their use of SATs would be particularly useful. One of the most valuable results of this research could be a catalogue of practical suggestions for analysts, based in part on the experience of practitioners and validated by careful examination of numerous IC products, on when to use SATs, how to employ them most effectively, and how to avoid problems that can undermine their analytic value.

ADDITIONAL APPROACHES

The IC might consider supplementary methods of assessing SATs. One possibility would be to examine the extent to which the use or nonuse of SATs correlates with measures of analytic quality as evaluated by the ODNI Office of Analytic Integrity and Standards and by agency-level evaluation programs. This approach could involve data mining to ascertain, at an aggregate level, which items employed SATs and how their evaluation scores compared with those of other pieces. A correlation

Broader research on SATs also would allow a greater focus on best practices and pitfalls in their use.

of pieces using SATs with higher or lower scores would not establish causality because other factors might have been the main influence on evaluators' ratings. But an SAT review of a selection of these pieces—along with others that did not employ SATs—would allow qualitative conclusions about the contribution of SATs, and a large sample would allow investigation of how particular SATs contributed or did not contribute to analytic quality.

A further element of future investigation could be controlled experiments testing the use of SATs. As noted previously, a major criticism of the few such experiments conducted so far in academia is that they did not replicate conditions in the IC. For example, they have generally tested the effect of SATs on individuals' thought processes, rather than on the IC group interactions that SATs are designed to facilitate.²⁷ A more-realistic approach would be to devise experiments based on current social science standards that compare the analytic results—measured by accuracy or by other criteria—of a group of analysts who use a specific SAT with those of a control group that does not use that SAT. Such an effort, although resource-intensive, could yield strong conclusions about the utility of specific analytic techniques:

- For example, senior CIA officials stressed to us the value of involving tradecraft methodologists and considering SATs at the onset of an analytic project. Further structured interviews and SAT reviews could assess how such efforts affect analytic quality.
- Additional study also could examine the potential utility of developing new SATs, particularly ones focusing on research design and use of big data.

Finally, a structured survey of analysts throughout the IC could provide otherwise unavailable data on the current use of SATs and provide a baseline for assessing their effectiveness in the future. Such a survey could provide statistically meaningful conclusions about the extent to which analysts are employing

SATs, which techniques are most widespread, and how analysts assess their value. The survey also could address why analysts are or are not using SATs.

One model for such an IC-wide effort could be the detailed tradecraft survey RAND recently completed with the U.S. Marine Corps Center for Marine Expeditionary Intelligence Knowledge. The survey examines the extent of analytic tradecraft practices in several areas, including structured techniques, factors that best explain the use of such techniques, and the reasons some respondents are not using established tradecraft. An IC-wide survey could cover these topics and explore the perceived utility of SATs in general and of specific SATs.

A complementary survey could examine the reaction to SAT-based analysis of intelligence consumers and their briefers. Insights from the latter group could be particularly valuable in addressing best practices and pitfalls because briefers have unparalleled insight into the needs and preferences of consumers but also understand the tradecraft of intelligence analysis.

Notes

- ¹ U.S. Government, "A Tradecraft Primer, Structured Analytic Techniques for Improving Intelligence Analysis," March 2009, p. 5. As of August 26, 2016: <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Tradecraft%20Primer-apr09.pdf>
- ² Intelligence Community Directive (ICD) 203, *Analytic Standards*, June 21, 2007.
- ³ Jack Davis, "Why Bad Things Happen to Good Analysts," in Roger Z. George and James B. Bruce, *Analyzing Intelligence: National Security and Practitioners' Perspectives*, 2nd ed., Washington, D.C.: Georgetown University Press, 2014, p. 131.
- ⁴ Interview with DIA officer, July 22, 2014.
- ⁵ Vincent Stewart, Drew Cukor, Joseph Larson, and Matthew Pottinger, "New Analytic Techniques for Tactical Military Intelligence," in George and Bruce, 2014, pp. 249–264.
- ⁶ Office of the Director of National Intelligence, *Analytic Transformation: Unleashing the Potential of a Community of Analysts*, McLean, Va., September 1, 2008, p. 17.
- ⁷ Office of the Director of National Intelligence, *The National Intelligence Strategy of the United States of America*, 2014, p. 7.
- ⁸ Randolph H. Pherson, "The Five Habits of the Master Thinker," *Journal of Strategic Security*, Vol. 6, No. 3, Fall 2013.
- ⁹ Richards J. Heuer, Jr., and Randolph H. Pherson, *Structured Analytic Techniques for Intelligence Analysis*, 1st ed., Washington, D.C.: CQ Press, 2011, p. 309.
- ¹⁰ Sundri Khalsa, "The Intelligence Community Debate over Intuition Versus Structured Technique: Implications for Improving Intelligence Warning and Analysis," *Journal of Conflict Studies*, Vol. 29, 2009. As of June 7, 2016: <http://journals.hil.unb.ca/index.php/jcs/article/view/15234/20838#no122>.
- ¹¹ Tim Van Gelder, "Can We Do Better Than ACH?" *AIPIO News*, December 2008.
- ¹² Randolph H. Pherson and Richards J. Heuer, "Structured Analytic Techniques: A New Approach to Analysis," in George and Bruce, 2014, pp. 231–235.
- ¹³ Jason U. Manosevitz, "Needed: More Thinking About Conceptual Frameworks for Analysis—The Case of Influence," *Studies in Intelligence* Vol. 57, No. 4, Extracts, December 2013.
- ¹⁴ Manosevitz, 2013; Gary H. McClelland, "Use of Signal Detection Theory as a Tool for Enhancing Performance and Evaluating Tradecraft in Intelligence Analysis," in Baruch Fischhoff and Cherie Chauvin, *Intelligence Analysis: Behavioral and Social Scientific Foundations*, Washington, D.C.: National Academies Press, 2011, pp. 95–97; Stephen Marrin, "Intelligence Analysis: Structured Methods or Intuition?" *American Intelligence Journal*, Summer 2007, pp. 7–16.
- ¹⁵ Richards J. Heuer, Jr., "The Evolution of Structured Analytic Techniques," presentation to the National Academy of Science, National Research Council Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security, Washington, D.C., December 8, 2009. As of June 7, 2016: <http://www.scribd.com/doc/24729291/The-Evolution-of-Structured-Analytic-Techniques-Richards-Heuer-8-DEC-2009>
- ¹⁶ Robert D. Folker, *Intelligence Analysis in Theater Joint Intelligence Centers: An Experiment in Applying Structured Methods*, Joint Military Intelligence College, January 2000.
- ¹⁷ Brant A. Cheikes, Mark J. Brown, Paul E. Lehner, and Leonard Adelman, *Confirmation Bias in Complex Analyses*, Bedford, Mass.: Mitre Corporation, 2004.
- ¹⁸ Philip E. Tetlock, *Expert Political Judgment*, Princeton, N.J.: Princeton University Press, 2005, pp. 190–202.
- ¹⁹ Heuer and Pherson, 2011, p. 313.
- ²⁰ Charlan Nemeth, Keith Brown, and John Rogers, "Devil's Advocate Versus Authentic Dissent: Stimulating Quantity and Quality," *European Journal of Social Psychology*, Vol. 31, 2001, pp. 707–720; Steven Rieber and Neil Thomason, "Creation of a National Institute for Analytic Methods: Toward Improving Intelligence Analysis," *Studies in Intelligence*, Vol. 49, No. 4, 2007. As of June 7, 2016: https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol49no4/Analytic_Methods_7.htm
- ²¹ Khalsa, 2009.
- ²² Philip E. Tetlock and Barbara A. Mellers, "Structuring Accountability Systems in Organizations: Key Trade-Offs and Critical Unknowns," in Fischhoff and Chauvin, 2011, pp. 253–257.
- ²³ Stephen Marrin, "Evaluating the Quality of Intelligence Analysis: By What (Mis) Measure?" *Intelligence and National Security* Vol. 27, No. 6, December 2012, pp. 325–339.
- ²⁴ Heuer, 2009.
- ²⁵ Marrin, 2012; Abbot E. Smith, "On the Accuracy of National Intelligence Estimates," *Studies in Intelligence*, Vol. 13, No. 4, 1994. As of June 7, 2016: https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/vol13no4/html/v13i4a04p_0001.htm
- ²⁶ Gregory T. Treverton, *Addressing "Complexities" in Homeland Security*, Sweden: Center for Asymmetric Threat Studies, 2009; Gregory T. Treverton, ed., *Comparing Early Warning Across Domains*, Sweden: Center for Asymmetric Threat Studies, 2011, p. 33.
- ²⁷ Richards J. Heuer, *Small Group Processes for Intelligence Analysis*, Monterey, Calif.: RJH Research, July 28, 2008. As of June 7, 2016: <http://www.pherson.org/wp-content/uploads/2013/06/02.-Small-Group-Processes.pdf>; Heuer and Pherson, 2011, pp. 312–316.

About This Report

Structured analytic techniques (SATs) are a key part of the rigorous analytic tradecraft that the Intelligence Community (IC) has pursued in recent years, but so far these techniques have received little systematic evaluation. This report argues that rigorous assessment of SATs is essential, albeit difficult; suggests specific questions that should be part of an SAT assessment; and proposes several methods the IC could consider to ascertain the practical value of SATs. The report also offers a pilot study that explores how larger-scale evaluation could be conducted via structured interviews with analytic practitioners and through a trial SAT review—a qualitative assessment of a body of IC products to examine the incidence and utility of SATs. These preliminary efforts are not intended to offer definitive conclusions about the value of SATs but rather to illustrate how the IC might evaluate them more systematically.

This research was conducted within the Intelligence Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the RAND Intelligence Policy Center, see www.rand.org/nsrd/ndri/centers/intel or contact the director (contact information is provided on the web page).

About The Authors

Stephen Artner is a RAND adjunct fellow who specializes in analytic methodology, European affairs, and political Islam.

Rich Girven is a senior international defense researcher and associate director of RAND's Intelligence Policy Center, who came to RAND from the Senate Select Committee on Intelligence, where he served as the Director of Analysis.

James Bruce, a Senior Political Scientist at the RAND Corporation, previously served nearly 24 years at CIA, chiefly as an analyst; he coedited *Analyzing Intelligence*, 2nd ed. (Georgetown University Press, 2014), and has taught graduate courses on intelligence at Georgetown University for the past 21 years.

Acknowledgments

The authors would like to thank Roger George and Vaughn Bishop for their thoughtful reviews of an earlier draft of this report. The analysis also benefited greatly from interviews with analytic methodologists and managers across the Intelligence Community.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.html.

For more information on this publication, visit www.rand.org/t/rr1408.

© Copyright 2016 RAND Corporation

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. **RAND®** is a registered trademark.

www.rand.org



NATIONAL DEFENSE RESEARCH INSTITUTE

RR-1408-OSD